

## Step 1 (Initialization)

Randomly choose T1 and T2 as the two cluster prototypes.

## Step 2

### Sub-step 2.1

$$J(T3, T1) = 1 / 7$$

$$J(T3, T2) = 3 / 5 = 0.6$$

So we associate T3 with T2 (we choose the one with large value in similarity because we want minimum distance)

$$J(T4, T1) = 3 / 7$$

$$J(T4, T2) = 5 / 5 = 1$$

So we associate T4 with T2

$$J(T5, T1) = 2 / 5 = 0.4$$

$$J(T5, T2) = 1 / 6$$

So we associate T5 with T1

$$J(T6, T1) = 4 / 5 = 0.8$$

$$J(T6, T2) = 2 / 7$$

So we associate T6 with T1

$$J(T7, T1) = 2 / 5 = 0.4$$

$$J(T7, T2) = 1 / 6$$

So we associate T7 with T1

$$J(T8, T1) = 4 / 7$$

$$J(T8, T2) = 4 / 7$$

So we associate T8 with T2 (so each cluster have the same size)

Cluster 1: T1, T5, T6, T7

Cluster 2: T2, T3, T4, T8

Sub-step 2.2

In cluster 1, for T1, the total similarity measure =  $0.4 + 0.8 + 0.4 = 1.6$

For T5,

$$J(T5, T6) = 1 / 5 = 0.2$$

$$J(T5, T7) = 0$$

The total similarity measure =  $0.4 + 0.2 = 0.6$

For T6

$$J(T6, T7) = 2 / 4 = 0.5$$

The total similarity measure =  $0.8 + 0.2 + 0.5 = 1.5$

For T7, the total similarity measure =  $0.4 + 0 + 0.5 = 0.9$

So we do not change the prototype for cluster 1

In cluster 2, for T2, the total similarity measure =  $0.6 + 1 + 4 / 7 = 2.17$

For T3

$$J(T3, T4) = 3 / 5 = 0.6$$

$$J(T3, T8) = 3 / 6 = 0.5$$

The total similarity measure =  $0.6 + 0.6 + 0.5 = 1.7$

For T4

$$J(T4, T8) = 4 / 7$$

The total similarity measure =  $1 + 0.6 + 4/7 = 2.17$

For T8

The total similarity measure =  $4/7 + 0.5 + 4/7 = 1.64$

So we do not change the prototype for cluster 2

So we do not need to repeat step 2

Step 3

**So the output is:**

**Cluster 1: T1, T5, T6, T7**

**Cluster 2: T2, T3, T4, T8**